

A. Baker



RAW SEQUENCE LISTING

DATE: 07/16/2002

PATENT APPLICATION: US/09/653,812

TIME: 14:22:27

Input Set : N:\Crf3\RULE60\09653812.raw

Output Set: N:\CRF3\07162002\I653812.raw

SEQUENCE LISTING

RECEIVED
JUL 19 2002
TECH CENTER 1600/2900

ENTERED

3 (1) GENERAL INFORMATION:

5 (i) APPLICANT: KAZAZIAN JR., HAIG H.
6 BOEKE, JEF D.
7 MORAN, JOHN V.
8 DOMBROSKI, BETH A.

10 (ii) TITLE OF INVENTION: COMPOSITIONS AND METHODS OF USE OF
11 MAMMALIAN RETROTRANSPOSONS

13 (iii) NUMBER OF SEQUENCES: 137

15 (iv) CORRESPONDENCE ADDRESS:

16 (A) ADDRESSEE: PANITCH SCHWARZE JACOBS & NADEL, P.C.
17 (B) STREET: ONE COMMERCE SQUARE, 2005 MARKET STREET, 22ND FL.
18 (C) CITY: PHILADELPHIA
19 (D) STATE: PA
20 (E) COUNTRY: U.S.A.
21 (F) ZIP: 19103-7086

23 (v) COMPUTER READABLE FORM:

24 (A) MEDIUM TYPE: Floppy disk
25 (B) COMPUTER: IBM PC compatible
26 (C) OPERATING SYSTEM: PC-DOS/MS-DOS
27 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30

29 (vi) CURRENT APPLICATION DATA:

C--> 30 (A) APPLICATION NUMBER: US/09/653,812
C--> 31 (B) FILING DATE: 01-Sep-2000
32 (C) CLASSIFICATION:

45 (vii) PRIOR APPLICATION DATA:

36 (A) APPLICATION NUMBER: US/08/847,844
37 (B) FILING DATE: 28-APR-1997
42 (A) APPLICATION NUMBER: US 08/749,805
43 (B) FILING DATE: 16-NOV-1996
46 (A) APPLICATION NUMBER: US 60/006,831
47 (B) FILING DATE: 16-NOV-1995

49 (viii) ATTORNEY/AGENT INFORMATION:

50 (A) NAME: DOYLE LEARY Ph.D., KATHRYN
51 (B) REGISTRATION NUMBER: 36,317
52 (C) REFERENCE/DOCKET NUMBER: 9596-23U2

54 (ix) TELECOMMUNICATION INFORMATION:

55 (A) TELEPHONE: 215-567-2020
56 (B) TELEFAX: 215-567-2991

59 (2) INFORMATION FOR SEQ ID NO: 1:

61 (i) SEQUENCE CHARACTERISTICS:

62 (A) LENGTH: 14 amino acids
63 (B) TYPE: amino acid

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/653,812

DATE: 07/16/2002

TIME: 14:22:27

Input Set : N:\Crf3\RULE60\09653812.raw

Output Set: N:\CRF3\07162002\I653812.raw

```

64      (C) STRANDEDNESS: single
65      (D) TOPOLOGY: linear
66      (ii) MOLECULE TYPE: protein
72      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
74      Met Val Gln Leu Lys Ile Leu Tyr Trp Asn Val Gly Lys Ser
75      1          5          10
77 (2) INFORMATION FOR SEQ ID NO: 2:
78      (i) SEQUENCE CHARACTERISTICS:
79          (A) LENGTH: 10 amino acids
80          (B) TYPE: amino acid
81          (C) STRANDEDNESS: single
82          (D) TOPOLOGY: linear
83      (ii) MOLECULE TYPE: protein
84      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
86      Tyr Asp Ile Val Ala Ile Gln Glu Pro Gly
87      1          5          10
89 (2) INFORMATION FOR SEQ ID NO: 3:
90      (i) SEQUENCE CHARACTERISTICS:
91          (A) LENGTH: 10 amino acids
92          (B) TYPE: amino acid
93          (C) STRANDEDNESS: single
94          (D) TOPOLOGY: linear
95      (ii) MOLECULE TYPE: protein
96      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
98      Lys Gly Arg Ala Val Ile Tyr Val Asn Lys
99      1          5          10
101 (2) INFORMATION FOR SEQ ID NO: 4:
102      (i) SEQUENCE CHARACTERISTICS:
103          (A) LENGTH: 13 amino acids
104          (B) TYPE: amino acid
105          (C) STRANDEDNESS: single
106          (D) TOPOLOGY: linear
107      (ii) MOLECULE TYPE: protein
108      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
110      Pro Thr Thr Val Tyr Ser Ile Tyr Ser Pro Ile Leu Thr
111      1          5          10
113 (2) INFORMATION FOR SEQ ID NO: 5:
114      (i) SEQUENCE CHARACTERISTICS:
115          (A) LENGTH: 16 amino acids
116          (B) TYPE: amino acid
117          (C) STRANDEDNESS: single
118          (D) TOPOLOGY: linear
119      (ii) MOLECULE TYPE: protein
120      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
122      Asn Leu Val Ala Val Gly Asp Leu Asn Leu His His Pro Asp Trp Asp
123      1          5          10          15
125 (2) INFORMATION FOR SEQ ID NO: 6:
126      (i) SEQUENCE CHARACTERISTICS:
127          (A) LENGTH: 24 amino acids

```

RAW SEQUENCE LISTING

DATE: 07/16/2002

PATENT APPLICATION: US/09/653,812

TIME: 14:22:27

Input Set : N:\Crf3\RULE60\09653812.raw

Output Set: N:\CRF3\07162002\I653812.raw

```

154         (B) TYPE: amino acid
155         (C) STRANDEDNESS: single
156         (D) TOPOLOGY: linear
158     (ii) MOLECULE TYPE: protein
163     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
165     Gly Glu Pro Thr Arg Leu Gly Asn Ala Thr Arg Gly Glu Arg Asp Gly
166         1             5             10             15
168     Thr Ile Asp His Ala Trp Leu Ser
169         20
171 (2) INFORMATION FOR SEQ ID NO: 7:
173     (i) SEQUENCE CHARACTERISTICS:
174         (A) LENGTH: 13 amino acids
175         (B) TYPE: amino acid
176         (C) STRANDEDNESS: single
177         (D) TOPOLOGY: linear
179     (ii) MOLECULE TYPE: protein
184     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:
186     Gly Ser Asp His Cys Pro Gln Glu Ile Trp Val Gln Val
187         1             5             10
189 (2) INFORMATION FOR SEQ ID NO: 8:
191     (i) SEQUENCE CHARACTERISTICS:
192         (A) LENGTH: 19 amino acids
193         (B) TYPE: amino acid
194         (C) STRANDEDNESS: single
195         (D) TOPOLOGY: linear
197     (ii) MOLECULE TYPE: protein
202     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:
204     Asp Ile Glu Gln Asn Pro Gly Pro Ile Ala Val Leu Gln Met Asn Val
205         1             5             10             15
207     Ser Cys Leu
210 (2) INFORMATION FOR SEQ ID NO: 9:
212     (i) SEQUENCE CHARACTERISTICS:
213         (A) LENGTH: 10 amino acids
214         (B) TYPE: amino acid
215         (C) STRANDEDNESS: single
216         (D) TOPOLOGY: linear
218     (ii) MOLECULE TYPE: protein
223     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:
225     Ala Asp Ile Ile Ala Ile Gln Glu Thr Trp
226         1             5             10
228 (2) INFORMATION FOR SEQ ID NO: 10:
230     (i) SEQUENCE CHARACTERISTICS:
231         (A) LENGTH: 10 amino acids
232         (B) TYPE: amino acid
233         (C) STRANDEDNESS: single
234         (D) TOPOLOGY: linear
236     (ii) MOLECULE TYPE: protein
241     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:
243     Gly Gly Gly Val Ala Val Leu Val Arg Lys

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/653,812

DATE: 07/16/2002

TIME: 14:22:27

Input Set : N:\Crf3\RULE60\09653812.raw

Output Set: N:\CRF3\07162002\I653812.raw

```

244      1          5          10
246 (2) INFORMATION FOR SEQ ID NO: 11:
248      (i) SEQUENCE CHARACTERISTICS:
249          (A) LENGTH: 13 amino acids
250          (B) TYPE: amino acid
251          (C) STRANDEDNESS: single
252          (D) TOPOLOGY: linear
254      (ii) MOLECULE TYPE: protein
259      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:
261      Asp Leu Ile Val Ala Ser Ala Tyr Met Arg Pro Pro Pro
262      1          5          10
264 (2) INFORMATION FOR SEQ ID NO: 12:
266      (i) SEQUENCE CHARACTERISTICS:
267          (A) LENGTH: 16 amino acids
268          (B) TYPE: amino acid
269          (C) STRANDEDNESS: single
270          (D) TOPOLOGY: linear
272      (ii) MOLECULE TYPE: protein
277      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:
279      Pro Leu Leu Leu Cys Gly Asp Phe Asn Met His His Pro Gln Trp Glu
280      1          5          10          15
283 (2) INFORMATION FOR SEQ ID NO: 13:
285      (i) SEQUENCE CHARACTERISTICS:
286          (A) LENGTH: 21 amino acids
287          (B) TYPE: amino acid
288          (C) STRANDEDNESS: single
289          (D) TOPOLOGY: linear
291      (ii) MOLECULE TYPE: protein
296      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:
298      Gly Glu Ile Thr Thr Ala Arg Gly Thr Arg Glu Arg Ser Cys Ile Asp
299      1          5          10          15
301      Leu Thr Trp Ser Lys
302      20
304 (2) INFORMATION FOR SEQ ID NO: 14:
306      (i) SEQUENCE CHARACTERISTICS:
307          (A) LENGTH: 13 amino acids
308          (B) TYPE: amino acid
309          (C) STRANDEDNESS: single
310          (D) TOPOLOGY: linear
312      (ii) MOLECULE TYPE: protein
317      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:
319      Leu Ser Asp His Tyr Val Leu Thr Phe Thr Leu His Gln
320      1          5          10
322 (2) INFORMATION FOR SEQ ID NO: 15:
324      (i) SEQUENCE CHARACTERISTICS:
325          (A) LENGTH: 17 amino acids
326          (B) TYPE: amino acid
327          (C) STRANDEDNESS: single
328          (D) TOPOLOGY: linear

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/653,812

DATE: 07/16/2002

TIME: 14:22:27

Input Set : N:\Crf3\RULE60\09653812.raw

Output Set: N:\CRF3\07162002\I653812.raw

```

330      (ii) MOLECULE TYPE: protein
335      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:
337      Met Asp Ile Arg Pro Arg Leu Pro Ile Gly Gln Ile Asn Leu Gly Gly
338      1          5          10          15
340      Ala
343 (2) INFORMATION FOR SEQ ID NO: 16:
345      (i) SEQUENCE CHARACTERISTICS:
346          (A) LENGTH: 10 amino acids
347          (B) TYPE: amino acid
348          (C) STRANDEDNESS: single
349          (D) TOPOLOGY: linear
351      (ii) MOLECULE TYPE: protein
356      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:
358      Leu Asp Ile Val Leu Val Gln Glu Gln Tyr
359      1          5          10
361 (2) INFORMATION FOR SEQ ID NO: 17:
363      (i) SEQUENCE CHARACTERISTICS:
364          (A) LENGTH: 10 amino acids
365          (B) TYPE: amino acid
366          (C) STRANDEDNESS: single
367          (D) TOPOLOGY: linear
369      (ii) MOLECULE TYPE: protein
374      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:
376      Lys Ala Gly Val Tyr Ile Arg Asn Arg Val
377      1          5          10
379 (2) INFORMATION FOR SEQ ID NO: 18:
381      (i) SEQUENCE CHARACTERISTICS:
382          (A) LENGTH: 13 amino acids
383          (B) TYPE: amino acid
384          (C) STRANDEDNESS: single
385          (D) TOPOLOGY: linear
387      (ii) MOLECULE TYPE: protein
392      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:
394      Asp Leu Tyr Met Val Ser Ala Tyr Phe Gln Tyr Ser Asp
395      1          5          10
397 (2) INFORMATION FOR SEQ ID NO: 19:
399      (i) SEQUENCE CHARACTERISTICS:
400          (A) LENGTH: 16 amino acids
401          (B) TYPE: amino acid
402          (C) STRANDEDNESS: single
403          (D) TOPOLOGY: linear
405      (ii) MOLECULE TYPE: protein
410      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:
412      Arg Val Val Ile Cys Ala Asp Thr Asn Ala His Ser Pro Leu Trp His
413      1          5          10          15
416 (2) INFORMATION FOR SEQ ID NO: 20:
418      (i) SEQUENCE CHARACTERISTICS:
419          (A) LENGTH: 21 amino acids
420          (B) TYPE: amino acid

```

VERIFICATION SUMMARY

DATE: 07/16/2002

PATENT APPLICATION: US/09/653,812

TIME: 14:22:28

Input Set : N:\Crf3\RULE60\09653812.raw

Output Set: N:\CRF3\07162002\I653812.raw

L:30 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:31 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:2556 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:137
L:2559 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:137